

# Trovare e valutare le informazioni online

*“Provare ad ottenere informazioni da Internet  
può essere come provare a bere un bicchiere  
d’acqua da un idrante”.*

Mitchell Kapor

## Introduzione

Il web oggi racchiude un numero enorme di contenuti, basti pensare che si stima esistano circa 1,3 miliardi di siti internet. Di fronte a questa vastità di dati, le sfide che ci troviamo davanti sono primariamente due: trovare le informazioni che cerchiamo e valutare la loro qualità e affidabilità. I contenuti del web, infatti, sono di diversa natura e distinguere le informazioni attendibili da quelle che non lo sono e le fonti affidabili da quelle non affidabili, non è semplice. Esistono strumenti in grado di assisterci nelle nostre ricerche, cioè i motori di ricerca, ma ad oggi nella selezione delle informazioni l’unico strumento a cui possiamo affidarci con solidità è la nostra capacità critica, che dobbiamo affinare seguendo alcune regole basilari.

## Trovare le informazioni attraverso i motori di ricerca

I motori di ricerca sono software progettati per cercare delle informazioni all’interno del web. Esistono due tipi di motori di ricerca, cioè motori di ricerca generici e motori di ricerca specializzati, che sono dedicati a tipologie di informazioni specifiche. Oggi il motore di ricerca generico più diffuso e conosciuto è sicuramente Google, ma è bene tenere a mente che esistono delle alternative (ad esempio Bing, DuckDuckGo, Ecosia, Yahoo), sebbene esse siano meno usate. I motori di ricerca specializzati, invece, servono a individuare e comparare informazioni specifiche, come il costo dei voli aerei in diverse compagnie di volo, o il prezzo dei prodotti in vendita in diversi negozi online (ad esempio Trovaprezzi, Skyscanner).

## Come cercare le informazioni sul web

La prima domanda che è necessario porsi quando si cerca un’informazione in Internet è: di che tipo di informazione si tratta? A seconda del tipo di informazione, infatti, cambia anche lo strumento di ricerca che è necessario impiegare. Se sto cercando un articolo scientifico, per esempio, sarà meglio non usare Google ma Google Scholar (<https://scholar.google.it/>), un motore di ricerca specificamente dedicato alle pubblicazioni accademiche. Se voglio invece trovare informazioni che riguardano il curriculum vitae di una persona, invece, posso cercare il suo nome all’interno di un social network professionale (ad es. LinkedIn, o Researchgate). Ma anche dopo una semplice ricerca con Google, possiamo notare come l’interfaccia che compare al di sopra dei risultati offra diverse opzioni (Tutti, Notizie, Immagini, Video, ecc.). Selezionare il tipo di risorsa informativa più adeguata, spesso ci permette di velocizzare molto le nostre ricerche e concentrarci sui risultati più pertinenti. Per esempio, se stiamo cercando informazioni che riguardano un fatto di cronaca recentemente avvenuto, selezionando l’opzione “Notizie” avremo più probabilità di individuare velocemente gli articoli dei quotidiani online che ne hanno parlato.

Un’altra opzione che è bene conoscere è la cosiddetta “ricerca inversa”. Questa tipologia di ricerca avviene a partire da un file che già abbiamo e viene svolta per cercare risultati simili o per trovare la fonte in cui è presente quel file. In pratica, mentre di solito cerchiamo delle pagine web

che trattano un determinato argomento, facendo una ricerca inversa carichiamo un file dal nostro computer e lo confrontiamo con quelli presenti online. La ricerca inversa più diffusa, oggi, è quella delle immagini. Infatti, nella sezione “Immagini” di un motore come Google, cliccando sull’icona della macchina fotografica è possibile fare un upload di un’immagine dal nostro PC o Smartphone e cercare se essa è presente in qualche pagina web ed eventualmente capire qual è. Questa operazione può essere effettuata a partire da una immagine scaricata da un articolo su cui abbiamo dei dubbi, per verificare chi altro ha pubblicato la fotografia. Se la troviamo in un articolo datato o diverso da quello che ho letto, è probabile che la foto sia stata riciclata e/o decontestualizzata.

### Valutare le informazioni dei motori di ricerca

I motori di ricerca sono dei software, che funzionano in base ad algoritmi specifici (un algoritmo può essere definito come “un procedimento di calcolo descrivibile con un numero finito di regole che conduce al risultato dopo un numero finito di operazioni” <http://www.treccani.it/vocabolario/algoritmo/> ). L’algoritmo di Google si chiama Page Rank e funziona usando diversi criteri diversi (che costituiscono una sorta di “ricetta segreta” di Google). Tra di essi, tuttavia, quello che ha un peso maggiore è la misura di

«quante altre pagine contengono link a ogni pagina identificata in una ricerca. Fra tutte le pagine web in cui sono presenti le parole indicate nella richiesta, quelle con il maggior numero di link in ingresso provenienti da altri siti hanno un punteggio più alto nella graduatoria dei risultati e i link dai siti che sono a loro volta collegati a un gran numero di siti contano di più rispetto a quelli provenienti da siti meno popolari” (Rheingold 2012, p. 125)

Si tratta, sostanzialmente, di una misura di popolarità, ma la popolarità non sempre è sinonimo di qualità, anche se spesso, per fortuna, si accompagna ad essa. Per questa ragione, quando usiamo un motore di ricerca è necessario fare attenzione a non ritenere che la presenza di un risultato in cima alla lista costituisca automaticamente un indice di una sua maggior qualità e affidabilità rispetto agli altri. Certo, il fatto che un sito sia popolare e che compaia nella prima pagina è un buon segno, che però non dovrebbe essere sopravvalutato e interpretato in modo errato. A volte può accadere che un risultato presente nella seconda pagina dei risultati sia più rilevante per i nostri interessi e di maggior qualità rispetto ai risultati della prima pagina. L’algoritmo dei principali motori di ricerca generici viene aggiornato periodicamente. Attualmente sono in corso delle sperimentazioni per riuscire ad introdurre delle modifiche in grado di eliminare dai risultati le notizie meno attendibili.

### Wikipedia

Wikipedia è una enciclopedia online libera e aperta, cioè un’enciclopedia che non è stata scritta da professionisti ma dagli stessi utenti del web. Chiunque può accedere e modificare il contenuto di una pagina, se la ritiene poco esaustiva o se individua informazioni che gli paiono poco accurate. Questo tipo di enciclopedia, dunque, esiste e funziona principalmente grazie allo spirito cooperativo della “comunità” di internauti, che mettono in comune il loro sapere con l’intento di permettere agli altri di accedere alle loro informazioni. Per questo motivo, possiamo dire che essa nasce da intenti molto nobili e che è un grande progetto di condivisione del sapere. Tuttavia, Wikipedia presenta anche alcuni problemi di affidabilità. Non essendoci dipendenti che lavorano propriamente per questa piattaforma, il controllo della veridicità delle affermazioni è affidato alla buona volontà degli utenti e non a professionisti, e questo la rende più vulnerabile alle manipolazioni esterne. Il controllo degli utenti alle modifiche che vengono introdotte, infatti, non sempre avviene in modo sistematico ed è difficile da gestire perché pochi utenti controllano molte pagine (Goodwin, 2013). Se, per

esempio, consultiamo una pagina immediatamente dopo che qualcuno l'ha modificata inserendo informazioni non veritiere, leggeremo delle informazioni non affidabili e dovremo essere in grado di accorgercene, ma non sempre è facile farlo.

Un altro aspetto potenzialmente critico del progetto Wikipedia è legato al concetto di "sapere cooperativo" su cui essa si basa, cioè l'idea secondo cui mettendo assieme le nostre conoscenze possiamo creare una voce enciclopedica esaustiva (si veda Goodwin, 2013). Questo sistema sembra funzionare bene nel caso di alcuni tipi di argomenti, pensiamo ad esempio alla creazione di una voce sulla storia di una squadra di calcio. In questo caso, raccogliendo le informazioni possedute da persone diverse possiamo costruire una risorsa conoscitiva molto articolata e ricca. Tuttavia, esistono argomenti più complessi che è più difficile trattare in modo esclusivamente cooperativo. Pensiamo a una voce enciclopedica su un poeta, come Petrarca. Immaginiamo che la voce di Wikipedia sia stata creata da centinaia di utenti che lo hanno studiato al liceo o al massimo ad un corso di letteratura medievale all'università, e confrontiamola ipoteticamente con una voce di una enciclopedia tradizionale che è stata creata da un ricercatore che ha studiato tutta la vita il poeta e che è stato retribuito per dedicarsi alla stesura di quella voce. A quale delle due ci affideremmo? Molti di noi propenderebbero per la seconda opzione, che però ha lo svantaggio di essere a pagamento o di richiedere un nostro sforzo (se, per esempio, per consultare l'enciclopedia dovessimo recarci in biblioteca).

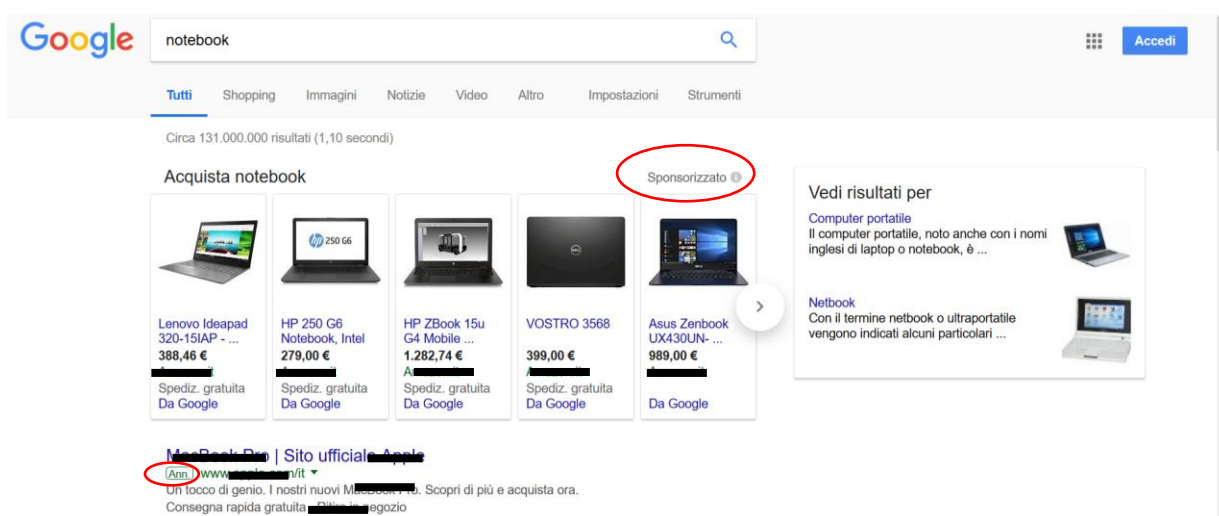
È interessante notare come tali aspetti critici (manipolabilità dei contenuti e limiti dell'approccio cooperativo) abbiano spinto Larry Sanger, uno dei fondatori di Wikipedia, ad abbandonare il progetto nel 2002 per dedicarsi a Citizendium, un'enciclopedia online che utilizza sia la partecipazione degli utenti sia quella degli esperti (che però oggi è diffusa molto meno rispetto a Wikipedia). L'esempio sopra descritto mette anche in luce in modo adeguato quanto detto in precedenza, cioè che quando cerchiamo delle informazioni dobbiamo sempre mettere a fuoco la tipologia di informazioni che stiamo cercando, quanto tempo abbiamo a disposizione (o vogliamo dedicare) per studiarlo e qual è l'uso che dobbiamo fare di quelle informazioni. Se abbiamo bisogno di farci un'idea indicativa rispetto a un autore o a un argomento, se abbiamo poco tempo a disposizione per la nostra ricerca e non dobbiamo riportare le informazioni a terzi (attraverso un altro articolo, ad esempio), allora Wikipedia è una risorsa indispensabile e utilissima. Se invece dobbiamo scrivere un articolo o una ricerca scolastica, se l'argomento è complesso oppure vogliamo studiare qualcosa in modo approfondito, allora probabilmente è meglio integrare la nostra ricerca con altre fonti informative.

### I risultati sponsorizzati

I motori di ricerca e i social network sono creati e sviluppati da aziende, che quindi devono trarre un profitto dal loro lavoro. I due sistemi principali attraverso cui entrambi "monetizzano" le nostre ricerche, cioè traggono guadagni, sono due. Il primo è la vendita di spazi pubblicitari ad aziende terze, che vengono inseriti tra i risultati dei motori di ricerca, il secondo è la vendita di dati sul comportamento degli utenti alle grandi aziende (che sono fortemente interessate a capire il comportamento dei consumatori per elaborare delle strategie in grado di massimizzare le vendite).

Di conseguenza, i contenuti che motori di ricerca e social ci offrono sono classificabili in due grandi categorie: risultati standard e risultati sponsorizzati. Questi ultimi sono annunci o contenuti che le aziende pagano per mostrare agli utenti, mentre i risultati standard sono le pagine che l'algoritmo ha individuato come rilevanti (nel caso dei motori di ricerca) o i contenuti (post, immagini, video) che l'algoritmo alla base del *newsfeed* dei social ha deciso di includere (magari perché, basandosi sui nostri *like* e commenti precedenti, crede che possano piacerci). Una ricerca condotta dall'Università Stanford, ha mostrato che i ragazzi americani (ma i risultati del nostro test indicano che anche per gli italiani le cose non sono molto diverse) spesso non sono in grado di

distinguere queste due tipologie di contenuti, cioè contenuti standard e sponsorizzati, in particolare quando usano i motori di ricerca (<https://www.internazionale.it/notizie/ananya-bhattacharya/2016/12/01/verificare-notizie-facebook-twitter>). Ovviamente, ciò li rende potenzialmente molto vulnerabili alle influenze delle aziende commerciali e poco consapevoli delle loro scelte online. Ciò accade anche perché all'interno dei *newsfeed* dei social (cioè le bacheche dove si scorrono i contenuti più recenti e rilevanti) le inserzioni sponsorizzate appaiono come molto simili ai contenuti standard e talvolta il loro riconoscimento è difficile. Nel caso dei motori di ricerca, gli inserzionisti sfruttano la nostra tendenza a concentrarci sui primi risultati che appaiono (cioè quelli più in alto) contrassegnando i loro link con un piccolo rettangolino verde al cui interno compare la scritta "Ann." o ponendo sopra i banner la scritta "sponsorizzato".



### Le fake news

Le cosiddette *fake news* sono notizie false pubblicate e diffuse online per motivi economici, politici o religiosi. È bene notare che le informazioni presenti online non si dividono nettamente in due categorie, cioè informazioni vere e *fake news*, dato che è possibile trovare pagine web che diffondono anche notizie imprecise o parzialmente vere. Mediamente la qualità del materiale presente nel web, purtroppo, non è eccellente. Per questo motivo è necessario individuare dei criteri attraverso cui valutare le informazioni. Basandoci sulle indicazioni che alcune biblioteche universitarie internazionali hanno stilato e che sono presenti online (ad esempio in questa pagina <http://guides.lib.berkeley.edu/evaluating-resources>) raccogliamo qui sotto alcuni criteri fondamentali che vengono usati nella valutazione delle informazioni. La soddisfazione di uno o due criteri, naturalmente, difficilmente sono sufficienti a contraddistinguere una informazione come affidabile. Maggiore è il numero di criteri che una pagina soddisfa, più alta è la probabilità che la notizia che stiamo leggendo è vera.

1. PRESENZA DI REVISIONE. Tipi di risorse diverse sono verificate in modo diverso: articoli scientifici, voci di enciclopedia aperta, voci di enciclopedia tradizionale, articoli giornalistici (iscritti all'albo), pagine di blog, pagine di siti personali. Mentre gli articoli scientifici delle riviste più prestigiose sono esaminati da figure autorevoli, che valutano la qualità del lavoro e decidono se pubblicarlo o meno, per pubblicare su un sito personale non è necessaria alcuna revisione.
2. ACCURATEZZA. Le informazioni devono essere descritte in modo accurato e non vago.
3. CORRETTEZZA LINGUISTICA e QUALITÀ DEI TESTI. Non devono essere presenti errori di ortografia o grammatica.
4. COLLEGAMENTO AD ALTRE RISORSE (considerate affidabili o autorevoli) E RIFERIMENTI BIBLIOGRAFICI.
5. AUTOREVOLEZZA. Alcuni siti web sono appartenenti ad enti che sono riconosciuti come autorevoli in alcuni campi o sono stati scritti da figure che riconosciamo come competenti in un ambito o su un argomento. Il sito del ministero della salute, uno scienziato conosciuto che ha delle pubblicazioni importanti, ecc.
6. ATTUALITÀ. Quanto le informazioni presenti nella pagina sono aggiornate e/o sono state scritte recentemente.
7. IMPARZIALITÀ. Quanto le informazioni sono descritte da un punto di vista parziale o imparziale.
8. USABILITÀ. Quanto il testo è facilmente leggibile e quanto è accompagnato da strumenti che facilitano la lettura e la comprensione, come indici, tabelle, grafici.

Esistono poi delle caratteristiche che spesso contraddistinguono una *fake news*. Uno dei primi a riconoscerli e ad elencarli è stato Craig Silverman, direttore di BuzzFeed.

- Il nome del sito riecheggia quello di quotidiani locali, ma è leggermente diverso.
- È accompagnata da una immagine che è stata usata in articoli che raccontano cose diverse.
- Riporta citazioni che in realtà sono legate ad altre notizie
- Non è presente la firma in calce al testo
- Usa titoli dal forte impatto emotivo, sensazionalistici
- Nessun sito autorevole o di quotidiani nazionali ha pubblicato quella informazione
- Indica un link ad una fonte più autorevole ma il link conduce solo alla home del sito (e non alla presunta notizia)

### La personalizzazione: perché la stessa ricerca fatta da due computer può dare risultati diversi

Alcune piattaforme come i motori di ricerca e i social network utilizzano oggi un sistema di personalizzazione dei contenuti. Ciò significa che esse cercano di registrare i nostri comportamenti online (cioè che tipo di siti abbiamo visitato, quali sono i nostri amici, quali post abbiamo commentato o a quali abbiamo messo *like*), per esempio attraverso i cosiddetti cookies, per poi mostrarci alcuni contenuti piuttosto che altri all'interno dei risultati dei motori di ricerca o dei *newsfeed* dei social network. L'intento è quello di aiutarci a trovare le informazioni che ci piacciono o ci interessano di più e di vendere spazi pubblicitari ritagliati ad hoc. I risultati sponsorizzati, di cui abbiamo parlato in precedenza, sono spesso personalizzati in base alle nostre ricerche e ai siti che abbiamo visitato (se recentemente abbiamo visitato siti di calzature, spesso troveremo inserzioni di venditori di calzature). Per l'utente, quindi, la personalizzazione rappresenta sia un vantaggio, in quanto velocizza le sue ricerche e rende il web più simile ai suoi gusti e alle sue preferenze, sia un pericolo, perché rende la pubblicità molto persuasiva ed efficace e ci chiude in una sorta di bolla

informazionale (Parisier 2012), le cui pareti riflettono i nostri pensieri rafforzando i *bias* e le distorsioni cognitive di cui parleremo nel prossimo paragrafo.

#### Lo studio del cervello: la ricerca di informazioni e la competizione per l'attenzione

Il fenomeno delle *fake news* è un fenomeno molto difficile da contrastare, per diverse ragioni, legate anche al nostro cervello e al suo funzionamento. Esistono, infatti, dei meccanismi cognitivi che sono sfruttati in modo molto efficace da chi diffonde queste informazioni. Uno di essi, è la nostra tendenza a interagire maggiormente con le informazioni che ci colpiscono emotivamente (Le Doux 1996). A causa di questa tendenza, siamo portati a cliccare, leggere e condividere più spesso notizie che hanno titoli che presentano un forte impatto emotivo, rispetto a quelle che vengono presentate con titoli più sobri ed equilibrati. Ovviamente, chi guadagna grazie ai click è portato a sfruttare tutto ciò, cercando in tutti i modi di colpire il lettore con un titolo che produce un effetto emotivo, che lo spinga a cliccare su quella notizia e – ancora meglio – a condividerla con le sue cerchie social. La tendenza a cliccare questo tipo di titoli, di conseguenza, è sfruttata tendenzialmente da chi crea *fake news* ma, potenzialmente, può esserlo da chiunque pubblica informazioni online. In un ambiente così ricco di stimoli come il web, infatti, si sta sviluppando una forte “competizione per l'attenzione” in cui le piattaforme e i siti, per riuscire a farci continuare a cliccare le loro pagine, in modo tale da garantire loro di guadagnare o semplicemente di sopravvivere, sono spinte a usare diversi espedienti. Tra questi espedienti, l'uso di titoli sensazionalistici ed emotivamente carichi è certamente uno dei più diffusi. Tuttavia, esso rischia di creare anche un circolo vizioso. Più clicchiamo su titoli emotivamente forti e sensazionali, più ci abitueremo a questo tipo di contenuti e più difficile sarà riuscire a catturare la nostra attenzione.

Un altro meccanismo che ostacola la corretta valutazione delle informazioni è chiamato *bias* di conferma (*confirmation bias*, in inglese). In psicologia, i *bias* sono errori sistematici che tendiamo a commettere quando ragioniamo. Un esempio classico di *bias* è la nostra tendenza a fare previsioni basandoci su quello che è successo in precedenza, anche quando non c'è alcun legame probabilistico tra due eventi. Se, per esempio, abbiamo tirato i dadi tre volte di fila e per tre volte è uscito il numero sei, saremo portati a credere che al quarto lancio la probabilità che esca di nuovo il numero sei sia minore, nonostante essa sia del tutto indipendente dai lanci precedenti, e quindi sempre e comunque uguale a 1/6.

Il *bias* di conferma è la nostra tendenza a cercare, leggere e credere alle informazioni che confermano le nostre credenze e a scartare quelle che le contraddicono (Nickerson 1998, Pasquinelli 2012). In pratica, quando cerchiamo informazioni, per natura (su questo gli psicologi cognitivi sono concordi) siamo propensi a cercare di confermare quello che crediamo e a scartare le pagine che riportano informazioni che contraddicono il nostro punto di vista. Per questo motivo, il *fact checking* – cioè quella pratica di smascheramento delle *fake news* attraverso la raccolta di informazioni attendibili che mostrano la falsità dell'informazione – non sembra oggi funzionare (Quattrocchi 2017) e non sembra rappresentare (finora) una soluzione alla loro diffusione. Un'altra conseguenza è che la personalizzazione, combinata con questa distorsione, risulta essere molto pericolosa. Immaginiamo, per esempio, che in base alle mie credenze io ritenga l'effetto serra una falsità e che in un social network io metta *like* ad alcune pagine che denunciano l'inesistenza dell'effetto serra. In base a questi apprezzamenti, in futuro il social network potrebbe inserire nel mio flusso di aggiornamenti (*newsfeed*) più spesso le notizie che negano l'effetto serra, cavalcando così la mia connaturata tendenza alla conferma. Potenzialmente, alla lunga ciò potrebbe rendere anche più *addictive* il social, cioè più in grado di generare dipendenza, perché ottenere conferme delle nostre credenze è gratificante, ma al contempo potrebbe danneggiare il nostro senso critico e distorcere la nostra visione della realtà (più del 90% degli scienziati che studiano il clima è concorde nel riconoscere l'effetto serra).

## Bibliografia

Goodwin, J. (2013), L'autorità di Wikipedia, in *Sistemi Intelligenti*, 1, pp. 9-37.

Gui, M. (2014), *A dieta di media. Comunicazione e qualità della vita*, Mulino.

LeDoux, J. (1996). *Il cervello emotivo. Alle origini delle emozioni*. Milano.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2), 175.

Pariser, E. (2012). *Il filtro. Quello che internet ci nasconde*. Il Saggiatore.

Pasquinelli, E. (2012), *Irresistibili Schermi. Fatti e misfatti della realtà virtuale*, Milano, Mondadori Università.

Quattrocioni, W. (2017). Inside the Echo Chamber. *Scientific American*, 316(4), 60-63.

Rheingold, H. (2012), *Perché la rete ci rende intelligenti*, Milano, Raffaello Cortina Editore.